Math 3339

Section 27204
MWF 10-11:00am AAAud 2

Bekki George
bekki@math.uh.edu
639 PGH

Office Hours:
M & Th noon – 1:00 pm & T 1:00 – 2:00 pm
and by appointment

Recall from 9/14:

Properties of Expected Value and Variance:

1. $E[c] = c$ for any constant $c \in \mathbb{R}$
2. $E[aX \pm bY] = aE[X] \pm bE[Y]$
3. $E[h(X)] = \sum_{x \in D} h(x) f(x)$ (*if discrete*); $E[h(X)] = \int_{-\infty}^{\infty} h(x) f(x)\, dx$ (*if continuous*)
4. $V(X) = E\left[(X-\mu)^2\right]$ or $V(X) = E[X^2] - E[X]^2$
5. $V(aX \pm b) = a^2 V(X)$

Note: Because of 5, $\sigma_{(aX+b)} = |a|\sigma_X$

Also, $V(aX+bY) = a^2V(X) + b^2V(Y)$ so, $\sigma_{(aX+bY)} = \sqrt{a^2V(X)+b^2V(Y)} \neq |a|\sigma_X + |b|\sigma_Y$

$a + b \neq \sqrt{a^2+b^2}$

The average stock price for companies making up the S&P 500 is $30, and the standard deviation is $8.20 (Business Week, Spring 2003). Assume the stock prices are normally distributed. $X \sim N(30, 8.2^2)$

a) What is the probability a company will have a stock price of at least $40?

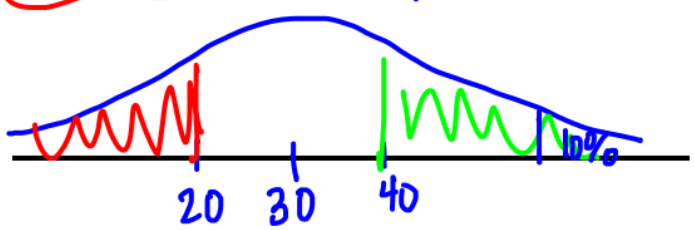
$P(X \geq 40) = 1 - P(X \leq 40) = 1 - \text{pnorm}(40, 30, 8.2)$

$1 - P\left(Z \leq \frac{40-30}{8.2}\right) = 1 - P(Z \leq 1.22)$

$= .1113 = 11.13\%$

b) What is the probability a company will have a stock price no higher than $20?

$P(X \leq 20) = \text{pnorm}(20, 30, 8.2) = .1113$



c) How high does a stock price have to be to put a company in the top 10%?

$P(X < \text{stock price}) = .9$

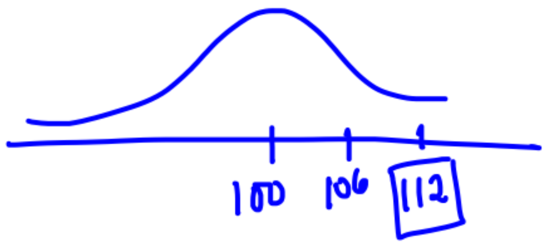
$\text{qnorm}(.9, 30, 8.2) = 40.5$

If a sample has a mean of 100 and a standard deviation of 6, what is the value in the set that corresponds to a z-score of 2?

$$\frac{X - \mu}{\sigma} = z\text{-score}$$

$$\frac{X - 100}{6} = 2$$

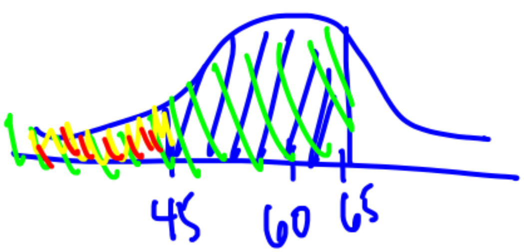
$$X = 112$$



The length of time needed to complete a certain test is normally distributed with a mean of 60 minutes and standard deviation of 10 minutes. What is the relative frequency of people who take between 45 and 65 minutes to complete the test?

$$P(45 \le X \le 65) = \text{pnorm}(65, 60, 10) - \text{pnorm}(45, 60, 10)$$

$$= .6247$$



$X \sim N(60, 10^2)$

Suppose there is another test (whose completion time is also normally distributed) with a mean of 45 minutes and a standard deviation of 7 minutes. What is the mean and variance time of someone who takes both tests back to back? $Y \sim N(45, 7^2)$

$Z = X + Y$

$\mu_Z = \mu_{X+Y} = 60 + 45 = 105$ min.

$\sigma_Z^2 = \sigma_{X+Y}^2 = 100 + 49 = 149$

$\sigma_Z = \sqrt{149} \approx 12.2$ min

Popper 13

1. If a sample has a mean of 48 and a standard deviation of 3.2, what is the value in the set that corresponds to a z-score of -1.2?
   a. 51.84
   b. 53.12
   c. 46.40
   d. 44.16
   e. none of these

2. Find $P(-1.9 < Z < 1.2)$
   a. 0.8020
   b. 0.7659
   c. 0.8562
   d. 0.9713
   e. none of these

$P(Z<1.2) = .8849$

$P(Z<-1.9) = .0287$

3. Find c such that $P(Z > c) = 0.8790$

a. -0.19

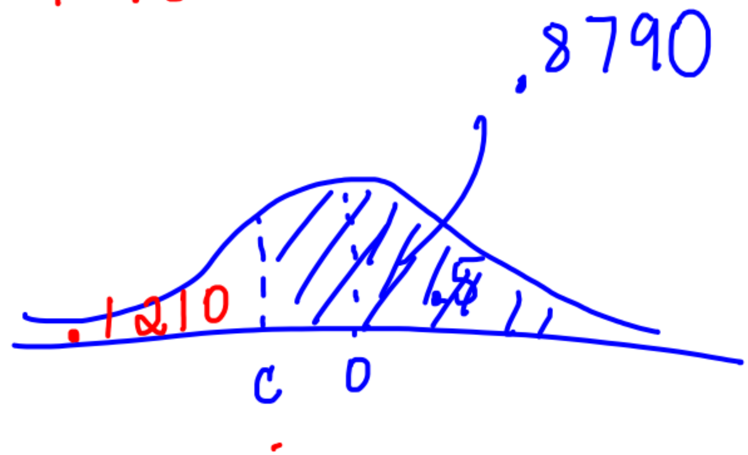
b. 1.17

c. -1.17

d. 0.19

e. none of these

$1 - .8790$



qnorm(.1210)

qnorm(1 - .8790)

Statistics and Their Distributions

$\mu =$ population mean

$\bar{X}$ = sample mean

Def: A ***statistic*** is any quantity whose value can be calculated from sample data.

σ^2 = pop. variance

s^2 = sample variance

Prior to obtaining the data, there is uncertainty as to what value of any particular statistic will result. Therefore, a statistic is a random variable and will be denoted by an uppercase letter. A lowercase letter represents the calculated or observed value of the statistic.

Def: The rv's $X_1, X_2, X_3, \ldots, X_n$ are said to form a (simple) ***random sample*** of size n if

1. The X_i's are independent random variables.
2. Every X_i has the same probability distribution.

Ex: A certain brand of MP3 player comes in three configurations:
2GB ($80), 4GB ($100), and 8GB ($120)

Let X = the cost of a single randomly selected purchase of the MP3 player. Suppose X has pmf given by the table below:

$E[X] = 80(.2) + 100(.3) + 120(.5) = 106$

x	80	100	120
$p(x)$	0.2	0.3	0.5

$V[X] = 11480 - 11236 = 244$

Suppose, on a particular day, 2 of these MP3 players are sold. Let

X_1 = selling price of first MP3 player

X_2 = selling price of second MP3 player

80, 80
80, 100
80, 120
100, 100
120, 120
100, 120

Determine the possible values for $\overline{X} = \dfrac{X_1 + X_2}{2}$

	80, 80	80, 100 or 100, 80			
$\overline{X}$	80	90	100	110	120
$p(\overline{x})$	.04	.12	.29	.30	.25

$E[\overline{X}] = 106$

$V[\overline{X}] = 122$

Give the pmf of $\overline{X} = \dfrac{X_1 + X_2}{2}$

Compute the expected value and variance of $\overline{X} = \dfrac{X_1 + X_2}{2}$

The Distribution of the Sample Mean

Suppose we take *n* measurements from a normal distribution with unknown mean. How can we approximate the mean?

Ex: Suppose the scores on Exam 1 are normally distributed and the average score is unknown. If 5 of the scores are

74 98 83 70 91

approximate the average score on the exam.

$$\bar{X} = 83.2$$

What about the spread? Will it be the same as the population of exams? NO

A simple example:

We will be looking at the "sampling distribution of the means" for a set of data – this is the same as the "distribution of the sample means".

Consider the population consisting of the values 3,5,9,11 and 14.

$\mu =$ 8.4 $\sigma =$ 4.4497

Let's take samples of size 2 from this population.

3,5 5,11
3,9 5,14
3,11 9,11
3,14 9,14
5,9 11,14

The set above is the **sampling distribution of size 2** for this population.

Let's take samples of size 3 from this population.

3, 5, 9

3, 9, 11

3, 9, 14 . . . . . .

The set above is the **sampling distribution of size 3** for this population.

List all the possible pairs/triples from 3,5,9,11 and 14 and find their means.

pairs	$\bar{x}$
3,5	4
3,9	6
3,11	7
3,14	8.5
5,9	7
5,11	8
5,14	9.5
9,11	10
9,14	11.5
11,14	12.5

$\mu_{\bar{x}} = \underline{8.4}, \sigma_{\bar{x}} = \underline{2.569}$

$\frac{6}{\sqrt{2}}$

triples	$\bar{x}$
3,5,9	17/3
3,5,11	19/3
3,5,14	22/3
⋮	⋮

$\mu_{\bar{x}} = \underline{8.4}, \sigma_{\bar{x}} = \underline{1.713}$

$\frac{6}{\sqrt{3}}$

Compare $\mu_{\bar{x}}$ (the mean of the sample means) to μ. same

What do you notice about $\sigma_{\bar{x}}$? $\frac{\sigma}{\sqrt{n}}$

Theorem:
If we sample from a normal distribution with mean μ and standard deviation σ, then the distribution of $\bar{X}$ is

$N\left(\text{mean} = \mu, \text{standard deviation} = \sigma / \sqrt{n}\right)$.

The Central Limit Theorem

If $\bar{X}$ is the mean of a random sample $X_1, X_2, \ldots, X_n$ from a distribution with mean μ and finite variance σ^2, then the distribution of

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

approaches the standard normal distribution as *n* approaches infinity.

$\boxed{n \geq 30}$

Popper 13

A waiter estimates that his average tip per table is \$20 with a standard deviation of \$4. If we take samples of 9 tables at a time, calculate the following probabilities when the tip per table is normally distributed.

μ ... $\sigma = 4$

4. What is the probability that his average tip is more than \$19?

   a. 0.599
   b. 0.773
   c. 0.707
   d. none of these

$P(\bar{x} > 19)$

$1 - P(\bar{x} \leq 19)$

$1 - \text{pnorm}(19, 20, 4/\sqrt{9})$

5. What is the probability that his average tip is between \$19 and \$22?

   a. 0.599
   b. 0.773
   c. 0.707
   d. none of these

$P(19 \leq \bar{x} \leq 22)$

$= \text{pnorm}(22, 20, 4/\sqrt{9})$

$- \text{pnorm}(19, 20, 4/\sqrt{9})$